

基于多元回归模型的大学生期末数学成绩 影响因素分析

刘云芬

(湖北师范大学 数学与统计学院 湖北 黄石 435002)

摘要: 对影响大学生期末数学成绩的相关因素进行了分析, 收集了各相关因素的数据, 建立了多元回归模型, 并基于 Matlab 软件对模型进行了求解分析. 结果表明: 学生高考成绩和平时成绩是影响大学生期末数学成绩的最主要的因素, 最后, 基于这两个因素, 对不同类型的大学生数学期末成绩进行了测算.

关键词: 多元回归; 高考成绩; 期末成绩; 平时成绩

中图分类号: O173.1 文献标识码: A 文章编号: 2096-3149(2018)04-0103-04

doi: 10.3969/j.issn.2096-3149.2018.04.020

0 引言

众所周知, 大学数学课程对高等学校培养创新型和应用型复合人才起着重要作用, 大学数学课程的学习有利于培养学生的思维素质. 虽然大学数学课程在各个高校广泛的开课, 但由于数学学科自身的特点, 学生普遍认为大学数学课程相对比较难学, 大学生的期末数学成绩也是越来越令人担忧. 对于此现象, 我们想要知道影响大学生数学期末成绩的主要因素是什么? 笔者结合自身的教学实践, 对湖北师范大学某班学生的数学成绩及影响因素进行了统计, 基于多元回归模型对以上问题进行一定的量化探讨.

1 多元线性回归^[1]

设自变量 x_1, x_2, \dots, x_p 的观测值 $x_{i1}, x_{i2}, \dots, x_{ip}$ 及因变量 y 对应的观测值 y_i 满足关系式:

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p$$

其中 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是相互独立且都服从正态分布 $N(0, \sigma^2)$ 的随机变量.

2 大学生期末数学成绩影响因素分析

2.1 大学生期末数学成绩影响因素

学生对知识的学习应该具有一定的延续性, 学生高考成绩或者是学生的高考数学成绩对大学数学的学习无疑有一定的影响, 文[2-4]探讨了大学生在校成绩同高考成绩之间的关系, 得出高考成绩对大学生在校数学成绩有较大影响的结论. 同时, 我们在教学实践中也看到, 有部分高考成绩相对较好的同学, 最后数学课程的期末考核不一定理想, 因此, 进入大学后的持续性学习态度也是影响大

收稿日期: 2018-04-12

基金项目: 湖北师范大学教研项目(2015028)

作者简介: 刘云芬(1979—), 女, 湖北鄂州人, 硕士, 讲师, 研究方向为不确定信息处理.

学数学成绩的一个重要因素。同时,由于大学生的学习相对于初高中的学习有很大的不同,大学生除了正常的课程学习之外,还要花大量的时间来参与课外实践活动,这一点在学生干部中更为突出,有少数学生干部课外实践活动做的非常出色,但是期末考试成绩却不太好。鉴于上面的分析,我们选取学生高考数学成绩、高考总成绩、是否班干部、大学期间的平时成绩作为大学数学成绩的影响因素。

2.2 数据来源

笔者对湖北师范大学某班同学进行了一个学期的观测,收集到该班同学的相关信息,在收集高考成绩时,剔除掉那些不是参加全国卷高考的同学信息;学生大学期间的平时成绩由学生的出勤率、作业情况以及期中检测综合而成,并将得到的数据进行归一化处理。相关数据见下表,其中,是否班干一列中,0表示非班干,1表示班干。

表1 某班成绩归一化数据表

序号	高考数学	高考总分	班干与否	平时成绩	期末成绩
1	0.8125	0.9917	0.0000	0.8947	0.7826
2	0.9375	0.9463	0.0000	1.000	0.9457
3	0.8571	0.9421	1.0000	0.8947	0.7717
4	0.7232	0.9421	0.0000	0.7368	0.5978
5	0.8571	0.9917	1.000	0.8947	0.7065
6	0.89294	0.9463	0.000	0.9474	0.8804
7	0.8400	0.9566	1.000	1.000	0.8804
8	0.7946	0.9628	1.000	0.9474	0.7717
9	0.8214	0.9566	0.000	1.000	0.8152
10	0.8393	0.9545	0.000	1.000	0.9130
11	0.8304	0.9483	1.000	0.7368	0.79351
12	0.8393	0.9442	1.000	0.8947	0.7826
13	0.8750	0.9793	0.000	0.8947	0.7826
14	0.7679	0.9401	0.000	0.9447	0.9130
15	0.8036	0.9421	0.000	1.000	0.9239
16	0.7946	0.9607	1.000	1.000	1.0000
17	0.8840	0.9545	0.000	1.000	0.8587
18	0.8571	0.9711	0.000	1.000	0.9130
19	0.9821	0.9588	0.000	1.000	0.9022
20	0.8482	0.9421	0.000	0.7368	0.7065
21	0.9375	0.9442	0.000	0.8947	0.8152
22	0.9643	0.9504	1.000	1.000	0.9130
23	1.0000	0.9690	1.000	0.7368	0.7065
24	0.7768	0.93801653	0.000	1.000	0.8804
25	0.8214	1.0000	0.000	1.000	0.8261

2.3 模型建立

基于上面的分析,记大学数学期末成绩为 y ,学生高考数学成绩为 x_1 、高考总成绩为 x_2 、是否班干部为 x_3 、大学期间的平时成绩为 x_4 ,以下将构建大学生数学期末成绩和诸影响因素之间的回归模型。

首先,假设大学生数学期末成绩和各因素之间是多元线性回归模型

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \quad (1)$$

利用 Matlab 统计工具箱进行求解,得到相关系数 $R^2 = 0.7134$,且回归系数的置信区间大部分都包含原点,显然结果不太满意。

进一步,对模型(1)进行改进,添加平方项,同时考虑是否班干部与平时成绩两因素的交互影响,进行残差分析并剔除异常数据后,得到如下回归模型:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1^2 + \beta_6 x_3^2 + \beta_7 x_4^2 + \beta_8 x_3 x_4 \quad (2)$$

仍然利用 Matlab 统计工具箱中进行求解,得到回归系数为:

$$\begin{aligned} \beta_0 &= -20.4576 & \beta_1 &= 42.0239 & \beta_2 &= 0.1775 & \beta_3 &= 6.5373 & \beta_4 &= -4.0452, \\ \beta_5 &= -3.5657 & \beta_6 &= -22.4703 & \beta_7 &= 2.8108 & \beta_8 &= -0.2183 \end{aligned}$$

同时,检验回归模型的三个统计量:相关系数 $R^2 = 0.9883$, $F = 105.7503$,与 F 对应的概率 $p = 0.0000$,并且回归系数的置信区间仅少数包含零点,说明模型回归比较理想。

2.4 模型结果分析

对于回归模型(2),求解大学生数学期末成绩对于四个因素偏导的最大值得到表 2:

表 2 四个因子对大学期末数学成绩的最大影响

是否班干部	高考数学成绩	高考总分	大学平时成绩
0.0166	1.3798	-0.1311	1.3731

由表 2 可得,高考数学成绩和大学期间平时成绩对大学期末数成绩影响相对较大,两者相差不是太大;是否班干部与高考总成绩对于大学期末数学成绩有影响,但是影响相对较小。

进一步,为了更好的解释模型,我们重点考虑高考数学成绩和大学平时成绩对大学数学期末成绩的影响。该班高考数学成绩最高分为 112 分,最低分为 81 分,将该班学生高考数学成绩分为三个档次,即是 100 分以上的记为 A 类,90 分到 100 分之间的记为 B 类,90 分以下的记为 C 类;大学期间平时成绩也分为三类,即是优秀(I)、一般(II)和较差(III)。基于回归模型(2),测算不同类型大学生的期末数学成绩,其中,由于该班高考总成绩的区分度不是太大,该因素取平均成绩进行测算,所得结果见表 3:

表 3 不同类的学生大学数学期末成绩测算

高考数学成绩	大学平时成绩	大学数学期末成绩
A 类	I	84.79 ~ 86.54
	II	75.00 ~ 77.26
	III	66.28 ~ 67.81
B 类	I	81.39 ~ 85.14
	II	73.59 ~ 75.86
	III	64.89 ~ 66.42
C 类	I	74.76 ~ 78.52
	II	66.96 ~ 69.23
	III	58.26 ~ 59.79

由表 3 可知,高考数学成绩属于 A 类或 B 类的同学,如果大学平时成绩属于 III 类,那么期末成绩是一个及格的水平;高考数学成绩属于 C 类的同学,如果大学平时成绩属于 I 类,那么期末成绩是可以达到中等水平的;高考数学成绩属于 C 类的同学,如果大学平时成绩属于 III 类,那么期末成绩则是不及格的水平。

3 结束语

通过收集影响大学生数学期末成绩的相关数据,建立了一个多元回归模型,基于该模型进行量化解释大学生期末数学成绩与其相关因素之间的关系,该模型总体比较理想,所得结果也比较符合实际,要想在大学数学期末考试中取得较好成绩,一定要端正学习态度,保质保量完成课程任务。

参考文献:

- [1]汪消银,周保平.数学建模与数学实验[M].北京:科学出版社,2015.
- [2]谢中才,郑惠娟.大学生高考成绩与大学阶段学习成绩的相关分析[J].数学的实践与认识,2009,39(12):1~6.
- [3]汪朝杰,谭常春,汪慧.大学生在校成绩与高考成绩的统计分析[J].大学数学,2013,29(4):80~86.
- [4]张丽华,戴学芳,刘志强,等.基于 Logistic 模型的大学生数学挂科原因实证分析[J].内蒙古师范大学学报(教育科学版),2014,27(9):136~138.

The influence factors for final mathematics achievement of college students based on multivariable regression

LIU Yun-fen

(College of Mathematics and Statistics , Hubei Normal University , Huangshi 435002 , China)

Abstract: Some influence factors related to final mathematics achievement in university were discussed. correlation factor data were collected , multivariable regression model was used to analyze the data. And the model was solved by Matlab software. the result showed that the mathematical scores in college entrance examination and usual performance were the uppermost factor , finally , the final mathematics achievement about different types college students were calculated based the two factor.

Key words: multivariable regression; grade of college entrance examination; final examination; usual performance